http://www.ajol.info/index.php/jrfwe
jfewr ©2018 - jfewr Publications
E-mail: jfewr@yahoo.com
ISBN: 2141 – 1778
Ogana, 2018

COMPARISON OF A MODIFIED LOG-LOGISTIC DISTRIBUTION WITH ESTABLISHED MODELS FOR TREE HEIGHT PREDICTION

Ogana F.N.
Department of Social and Environmental Forestry, University of Ibadan, Nigeria
E-mail: *ogana_fry@yahoo.com;* +2347037679328

ABSTRACT

The complexity of data structure from different forest stands across the world has necessitated the continuous introduction of new models in forestry. No single model is expected to provide accurate fit to all data sets. Therefore, in this study, the cumulative distribution function (cdf) of the Log-Logistic distribution was modified to construct a new height-diameter (h-d) model for Gmelina arborea Roxb plantation in Omo Forest Reserve, Nigeria. A total of 60 sample plots of 0.04 ha were used in this study. Tree diameters and heights measurement were taken on 1,189 trees. The new h-d model was termed M. LogL and its performance was compared with five established traditional h-d models that have been used in quantitative forestry study. These include: Logistic h-d, Chapman-Richards (C-R) h-d, Weibull h-d, Näslund h-d and Curtis h-d. Model assessment was based on adjusted R^2, root mean squared error (RMSE), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Hannan-Quinn Criterion (HQC). The result showed that the performance of the new M. LogL h-d was comparable to other traditional h-d models used in forestry. The adjusted R^2, RMSE, AIC, BIC and HQC were 0.629, 3.343, 2555.599, 2565.527 and 2555.480, respectively. However, the Logistic h-d model had the overall best fit to the data set. The order of ranking was: Logistic > M. LogL > Curtis > C-R > Weibull > Näslund. Therefore, the M. LogL model can be used to predict tree height for the Gmelina arborea plantation.

Keywords: Modified Log-Logistic, height-diameter, Gmelina arborea, Omo forest reserve.

INTRODUCTION

Tree height-diameter (h-d) models have consistently been used to estimate tree height for subsample of tree which diameter at breast height (dbh) has been measured. This is an attractive tool in forest inventory; because considerable time is required to measure tree height for large sample especially in natural forest. Also, direct tree height measurement for a large sample of trees is also expensive. Although sophisticated inventory equipment such as vertex II, range finder, among others, have been developed to reduce the time require to measure tree height. However, these instruments are not easily affordable. Therefore, the use of h-d model is inevitable so that for a given tree diameter the height can be estimated with some level of accuracies.

For any h-d relationship the following characteristics are desirable: monotomic increment, inflection point and have upper asymptotic value (Lei and Parresol, 2001). Some of the traditional h-d models that have been applied to forestry include: Logistic (Peer and Reed, 1920), Korf (Lundqvist, 1957), Chapman-Richards (Richards, 1959), Curtis (1967), Weibull (Yang *et al*., 1978), Wykoff *et al*. (1982), Ratkowsky (1990) etc. Several published forestry literatures on h-d models exist (Lynch and Murphy, 1995; López-Sanchez *et al*., 2003; Mehtätalo, 2005; Trincado *et al*., 2007; Temesgen *et al*., 2008; Krisnawati *et al*., 2010; Coble and Lee, 2011; Mehtätalo *et al*., 2015; Shamaki *et al*., 2016, Eby *et al*., 2017 and so on). Shamaki *et al.* (2016) developed height-diameter relationship for *Tectona grandis* plantation in Nimbia Forest Reserve, Nigeria using Chapman-Richards and Weibull models. The authors found the Chapman-Richards function to be the best model for the plantation based on pseudo coefficient of determination and root mean square error. However, Huang *et al*. (1992) noted the Chapman-Richards model approaches the upper asymptote too quickly even when there is a week relationship between the predictor and response variables.

The irregularities and complexity of data structure from different forest stands across the world has necessitated the continuous introduction of new models in forestry. No single model is expected to provide accurate fit to all data sets. There is need to develop a simple model that is biologically realistic. Cumulative distribution functions (cdf) with closed-form such as the Log-Logistic can be explored further to develop more h-d models for forestry study. In this study, the cdf of the Log-Logistic distribution was modified to construct a new h-d model for *Gmelina arborea* Roxb plantation in Omo Forest Reserve. Its performance was compared with other established h-d models in forestry.

MATERIALS AND METHODS

Data

The data for this study were obtained from the *Gmelina arborea* plantation in Omo Forest Reserve. Its lies between Latitudes 6°35′ and 7°05′N and Longitudes 4°19′ and 4°10′E of Ijebu-East, and North Local Government Areas of Ogun State, Nigeria (Chima *et al*., 2009 cited in Ogana *et al*., 2017). It occupies an area of 130,500 ha. A total of 60 sample plots of 0.04 ha across different age series were used. Tree diameters and heights measurement were taken on 1,189 trees. Ten percent (10%) of the data were used for model validation. The descriptive statistics of the inventory data is presented in Table 1.

Table 1: Descriptive statistics of the data set

	Statistics			
Stand variables	**Mean**	**Max**	**Min**	**Standard deviation**
Dbh (cm)	19.02	49.65	4.65	8.79
THt (m)	13.41	42.7	1.6	5.78
Stand age (yrs)	24	33	12	8.91
Quadratic mean diameter (cm)	20.14	26.83	11.18	3.67
Dominant height (m)	18.79	29.85	8.28	5.71
Density (N/ha)	495.42	825	250	152.9
Basal area (m^2/ha)	17.09	38.23	2.57	8.79
Volume (m^3/ha)	224.06	845.86	15.94	166.11

Model specification

The modified Log-Logistic (M. LogL) h-d model was derived from the cumulative distribution function (cdf) of the 2-parameter Log-Logistic distribution. The cdf of a random variable X say, gives the probability that the variable gets the value of x or less. The Log-Logistic distribution has a closed-formed cdf expressed as:

$$F(x) = \left[1 + \left(\frac{\alpha}{x}\right)^{\beta}\right]^{-1}$$

Where:

$F(x)$ = cumulative distribution function;

x = random variable;

, β = scale and shape parameters of the distribution.

Equation 1 was modified by the inclusion of the point of measurement of diameter at breast height (dbh) i.e., 1.3 m above the ground. Also, the constant plus 1 was removed from the equation to make the prediction more reasonable. The new M. LogL h-d is presented in Table 2. This M. LogL h-d was compared with five established tradition h-d models that have been used in quantitative forestry study. These include: Logistic h-d, Chapman-Richards (C-R) h-d, Weibull h-d, Näslund h-d and Curtis h-d. Altogether, six h-d models were evaluated in this study. To prevent the problems associated with back transformation bias, only nonlinear h-d models were considered (Table 2).

Table 2: Model specification

Name	Model	References	Eq.
M. LogL	$h = 1.3 + \left[\left(\frac{\alpha}{d}\right)^{\beta}\right]^{-1}$		(2)
Logistic	$h = 1.3 + \frac{\alpha}{1+\beta e^{-\gamma d}}$	Pearl and Reed (1920), Mehtätalo *et al*. (2015)	(3)
C-R	$h = 1.3 + \alpha\left(1 - e^{(-\beta d)^{\gamma}}\right)$	Richards (1959), Mehtätalo *et al*. (2015)	(4)
Weibull	$\mathrm{h} = 1.3 + \alpha\left(1 - e^{-\beta d^{\gamma}}\right)$	Weibull (1951), Mehtätalo *et al*. (2015)	(5)
Näslund	$h = 1.3 + \frac{d^2}{(\alpha d + \beta)^2}$	Näslund (1937), Mehtätalo *et al*. (2015)	(6)
Curtis	$h = 1.3 + \frac{(\alpha d)}{(1+d)^{\beta}}$	Curtis (1967), Mehtätalo *et al*. (2015)	(7)

h = height; d = diameter at breast height; α, β, γ = model parameters *e* = exponential; C-R = Chapman-Richards; M. LogL = modified Log-Logistic

The models were fitted using nonlinear regression procedure from *nlstools* package in R (R Core Team, 2026; Baty and Delignette-Muller, 2015). Model assessment was based on adjusted coefficient of determination (adj R^2), root mean square Error (RMSE), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Hannan-Quinn Criterion (HQC). Information criteria measure the difference between the models being assessed and the "true" model that is being sought. Models with high adj R^2 and low RMSE, AIC, BIC and HQC were regarded as good models

$$Adj\ R^2 = 1 - \frac{(n-1)}{(n-p)} \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2} \quad (8)$$

$$AIC = n \ln\left(\frac{RSS}{n}\right) + 2p \quad (9)$$

$$BIC = n \ln\left(\frac{RSS}{n}\right) + p \ln n \quad (10)$$

$$HQC = n \ln\left(\frac{RSS}{n}\right) + p \ln \ln n \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}} \quad (12)$$

Where: RSS = residual sum of square, n = sample size, p = number of parameters; $\bar{Y}_i$ = average tree height; Y_i is the observed value and $\hat{Y}_i$ is the theoretical value predicted by the model. The models were also validated using independent data set (i.e. data not used in calibrating the models). The predicted and observed values from each model were compared using paired sample t-test at 5% probability level.

RESULTS AND DISCUSSION

The result showed that the best h-d models for the *Gmelina arborea* data were the Logistic h-d and the M. LogL h-d models. The Logistic h-d model ranked first with highest adjusted R^2, lowest RMSE, AIC, BIC and HQC of 0.638, 3.302, 2530.353, 2545.245 and 2530.175, respectively. This was followed by the new M. LogL h-d model with 0.629, 3.343, 2555.599, 2565.527 and 2555.480, respectively (Table 3). The Curtis, C-R, Weibull, and Näslund h-d models also performed well based on the fit indices. In fact, the fit indices of Curtis h-d model were the same with the new M. LogL h-d model. Judging the performance of the models based on AIC and as a common rule of thumb, two models are indistinguishable if the difference (i.e. ΔAIC) of their AICs value is less than 2. The absolute difference in the AIC values of M. LogL h-d and C-R, Weibull and Näslund was greater than 10 (the smaller the AIC the better the model). This shows that the performance of M. LogL was different and better than C-R, Weibull and Näslund h-d models. Furthermore, the validation of the models with independent data set showed that the predicted values from Logistic and M. LogL were not significant from the observed tree height at 5% probability level. The predicted height values from Curtis, C-R, Weibull, and Näslund h-d models were significantly different from the observed tree height. The order of ranking was: Logistic > M. LogL >

Curtis > C-R > Weibull > Näslund.

Mehtätalo *et al*. (2015) reported RMSE values of 2.20 and 2.79 for *Eucalyptus globulus* and mixed Tropical species, respectively for Logistic h-d model. While Naslund had RMSE of 2.29 and 2.81, respectively, for the two species in their study. Also, Shamaki *et al*. (2016) reported R^2 and RMSE of 0.336 and 0.995 and 0.335 and 0.995 for Chapman-Richards and Weibull h-d models, respectively, for *Tectona grandis* in Nimbia Forest Reserve. There is no basis for comparing the fit indices values of the study at hand and others reported in forestry literatures. As the parameters of the models are location specific and more importantly heavily depends on the quality of data used in fitting the models.

Table 3: Fit statistics of the fitted h-d model

Model	**Parameters**			**Fit Statistics**					**Validation**	
									Pred. vs Obs. T-test	
	α	**β**	**γ**	**Adj R^2**	**RMSE**	**AIC**	**BIC**	**HQC**	**T-value**	**P-value**
M. LogL	0.977	0.844		0.629	3.343	2555.599	2565.527	2555.480	1.977	0.050
Logistic	27.691	6.791	0.087	0.638	3.302	2530.353	2545.245	2530.175	1.579	0.117
C-R	52.419	0.04	0.353	0.626	3.359	2566.527	2581.420	2566.350	2.241	0.027
Weibull	37.525	0.023	0.969	0.622	3.373	2575.381	2590.273	2575.203	3.261	0.002
Näslund	0.162	2.257		0.607	3.441	2616.571	2626.499	2616.452	3.682	0.001
Curtis	1.052	0.163		0.629	3.343	2555.599	2565.527	2555.480	1.985	0.049
Calibration data N=1056									Validation N= 117	

C-R = Chapman-Richards; Adj R^2 = adjusted coefficient determination; RMSE = root mean square error; AIC = Akaike Information; Criterion; BIC = Bayesian Information Criterion; HQC = Hannan Quinn Criterion

The height curves produced by the h-d models followed the observed tree height pattern (Fig. 1a, b, c, d, e and f). However, the s-shaped was more obvious in the Logistic h-d model. The Näslund h-d function flattened too quickly. The height curve produced by the new M. LogL are comparable to other functions considered this study. The parameters estimate of the M. LogL were low and have asymptotically significant t-statistics.

The analysis of the plots of residual and predicted values demonstrated that there was little or no systematic bias towards over-or underestimation of the tree total height (Fig. 2a and b). The assumption of ordinary nonlinear least squares regression was not violated in the models as the residual analysis showed almost homogenous variance (homoscedasticity) over the range of the predicted values and no systematic patterns. Only the residual plot of Logistic h-d and M. LogL h-d models were presented. Huang *et al*. (1992) asserted that "for a good h-d model, the asymptotic t-statistic for each coefficient should be significant, the MSE should be small and the residual plot should approximate homogeneous variance over the full range of predicted values". The M. LogL h-d model met these criteria relative to other established h-d functions evaluated in this study.

One major challenge with the fitting of 3-parameter h-d models (e.g. Logistic, Chapman-Richards, Ratkowsky, Weibull etc.) is convergence (Mehtätalo *et al*., 2015). At times convergence may not be achieved during model-fitting process. When this occurs, a common practice is to set one of the parameters to fixed value. Of course, this will invariably affect the resulting performance of the h-d model, unless adequate value is chosen. But this is not a problem with most 2-parameter h-d models e.g. Curtis, Meyer, Michaelis-Menten, Näslund, Wykoff, etc. The performance of M. LogL h-d was better compared to these 2-parameter h-d models such as Curtis and Näslund, especially, when prediction was made with the independent data set.

CONCLUSION

The choice of height-diameter model may be contingent on the ease of achieving convergence to a solution, the model properties and its biological realism. The M. LogL h-d tested in this study possessed these futures. Thus, the model can be used to predict tree height for the *Gmelina arborea* plantations in Omo Forest Reserve, Nigeria. The M. LogL may be tested on other species from different location.

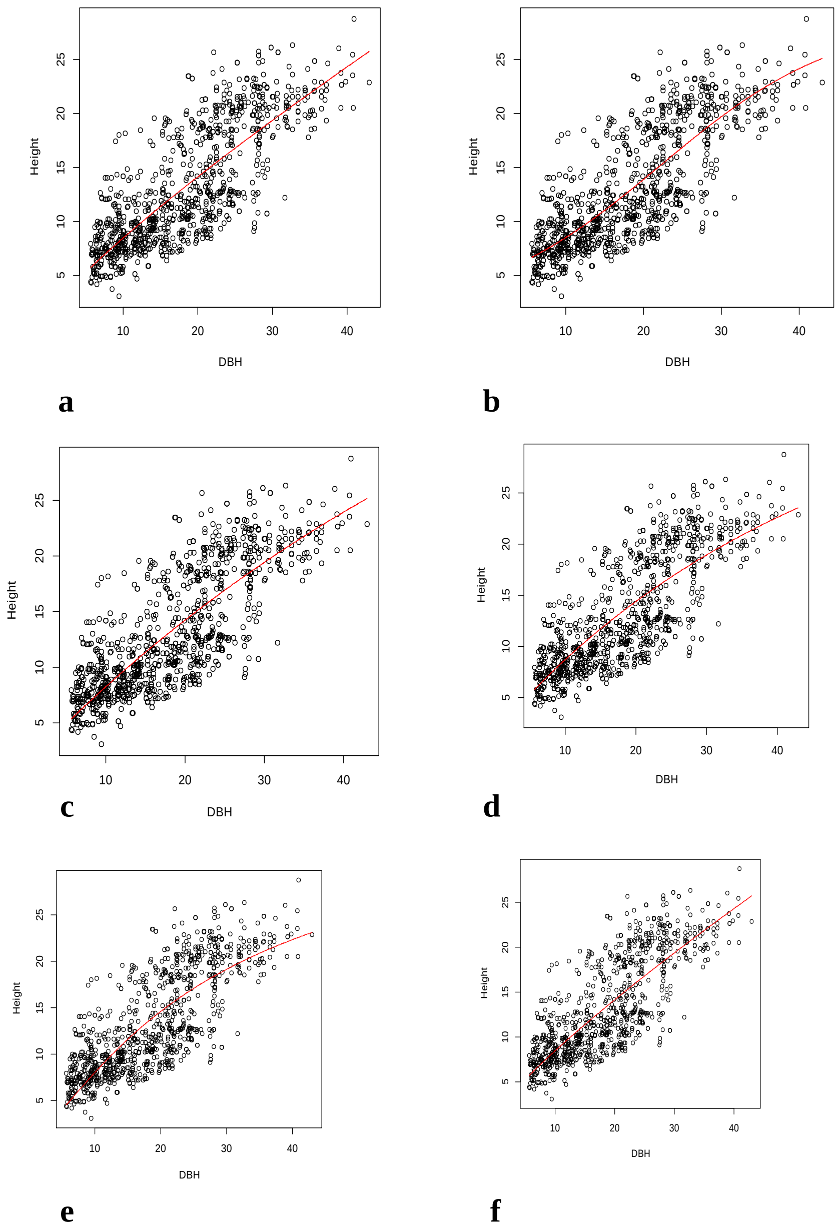

Fig. 1: Observed height (scatter dots) and fitted models (lines) for (a) M. LogL h-d, (b) Logistic h-d, (c) Chapman-Richards h-d, (d) Weibull h-d, (e) Näslund h-d and (f) Curtis h-d.

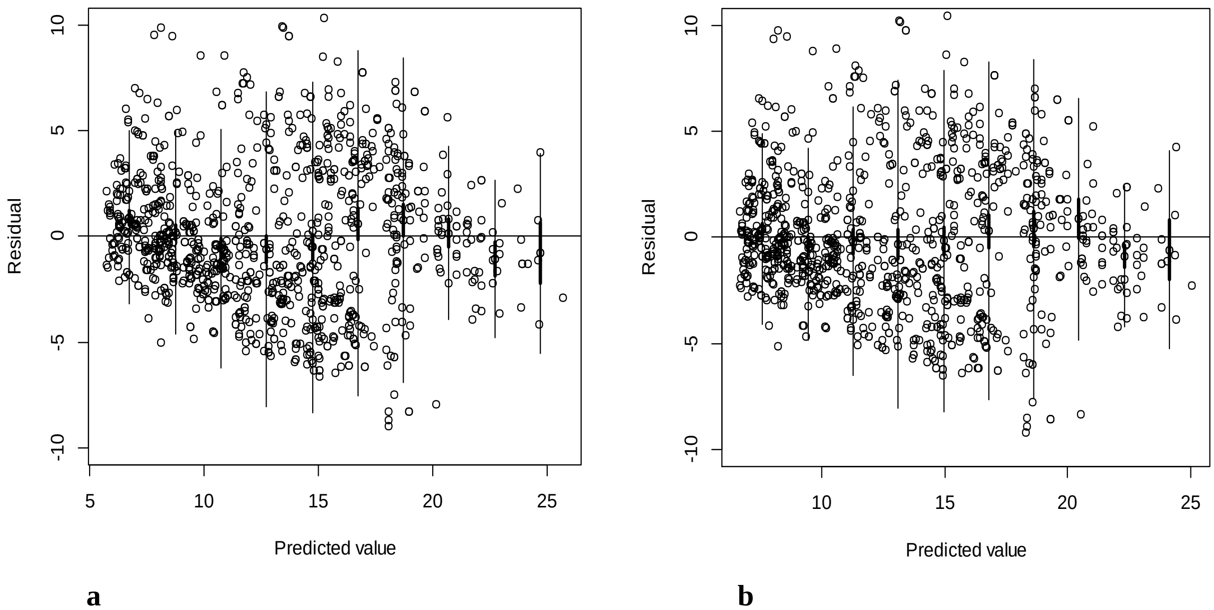

Fig. 2: Residual graphs for (a) M. LogL h-d and (b) Logistic h-d.

REFERENCES

Baty, F., & Delignetter-Muller, M. 2015. Tools for nonlinear regression analysis. R Package.

Chima, U.D., Ola-Adams, B.A., Adedire, M.O., & Oyelowo, O.J. 2009. Impact of Land Use Changes on Soil Quality of a Biosphere Reserve in Southwestern Nigeria. *Journal of Research in Forestry, Wildlife and Environment*, 1(1): 111-122

Curtis, R.O. 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. *Forest Science*, 13(4): 365-375.

Eby, W.M., Oyamakin, S.O. & Chukwu, A.U. 2017. A new nonlinear model applied to the height-DBH relationship in *Gmelina arborea*. *Forest. Ecology and Management.* 397:139-149

Huang, S., Titus, S.J. & Weins, D.P. 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. *Can. J. For. Res*. 22: 1297-1304.

Krisnawati, H., Wang, Y. & Ades, P.K. 2010. Generalized height-diameter models for *Acacia mangiun* Willd Plantations in South Sumatra. *Journal of Forestry Research*, 7(1): 1-19.

Lundqvist, B. 1957. On the height growth in cultivated stands for pine and spruce in Northern Sweden. *Medd. Frstatens skogforsk* 133

Lei, Y. & Parresol, B.R. 2001. Remarks on height-diameter modelling. *Res. Pap.* SRS-10. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 6p

López-Sanchez, C.A., Varela, J.G., Dorado, F.C., Aboreca, A.R., Soalleiro, R.R., Alvarez-Gonzalez, J.G. & Rodriguez, F.S. 2003. A height-diameter model for *Pinus radiate* D. Don in Galicia (Northwest Spain). *Ann. Forest Science*, 60:237-245.

Lynch, T.B. & Murphy, P.A. 1995. A compatible height prediction and projection system for individual trees in natural, even-aged shortleaf pine stands. *Forest Science*, 41: 194-209.

Mehtätalo, L., de-Miguel, S. & Gregoire, T.G. 2015. Modelling height-diameter curves for prediction. *Canadian Journal of Forest Research*, 45: 826-837

Mehtätalo, L. 2005. Height-diameter models for Scots pine and birch in Finland. *Silva Fenn*.

39(1): 55-66.

Näslund, M. 1937. Skogsförsöksanstaltens gallringsförsök I tallskog (Forest research institute's thinning experiments in Scots pine forests). Meddelanden frstatens skogsförsöksanstalt Häfte 29 (In Swedish).

Ogana, F.N., Itam, E.S. & Osho, J.S.A. 2017. Modelling diameter distributions of *Gmelina arborea* plantation in Omo Forest Reserve, Nigeria with Johnson's S_B. *Journal of Sustainable Forestry*, 36(2): 123-133

Pearl, R. & Reed, L.J. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proc. Natl. Acad. Sci. U.S.A. 6: 275–288.

R Core Team 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Ratkowsky, D.A. 1990. *Handbook of nonlinear regression*. Marcel Dekker, Inc., New York.

Richards, F.J. 1959. A flexible growth function for empirical use. *Journal of Experimental Biology*, 10: 290-300.

Shamaki, S.B., Akindele, S.O., Isah, A.D. & Mohammed, I. 2016. Height-diameter relationship models for Teak (*Tectona grandis*) plantation in Nimbia Forest Reserve, Nigeria. *Asian Journal of Environment and Ecology*, 1(1): 1-7.

Temesgen, H., Monleon, V.J. & Hann, D.W. 2008. Analysis and comparison of nonlinear tree height prediction strategies for douglas-fir forests. *Canadian Journal of Forest Research*, 38(3): 553-565.

Trincado, G., VanderSchaaf, C.L. & Burkhart, H.E. 2007. Regional mixed-effects height-diameter models for loblolly pine (*Pinus taeda* L.) plantations. *European Journal of Forest Research*, 126: 253-262

Wykoff, W.R., Crookston, N.L. & Stage, A.R. 1982. User's guide to the stand prognosis model. USDA For. Serv. Gen. Tech. Rep. INT-133

Yang, R.C., Kozak, A. & Smith, J.H.G. 1978. The potential of Weibull-type functions as a flexible growth curve. *Canadian Journal of Forest Research*, 8:424-431